

---

# **Investigación e Innovación Educativa en Docencia Universitaria.**

## **Retos, Propuestas y Acciones**

**Edición de.**

Rosabel Roig-Vila  
Josefa Eugenia Blasco Mira  
Asunción Lledó Carreres  
Neus Pellín Buades

**Prólogo de.**

José Francisco Torres Alfosea  
Vicerrector de Calidad e Innovación Educativa  
Universidad de Alicante

Edición de:

Rosabel Roig-Vila  
Josefa Eugenia Blasco Mira  
Asunción Lledó Carreres  
Neus Pellín Buades

© Del texto: los autores (2016)

© De esta edición:

Universidad de Alicante  
Vicerrectorado de Calidad e Innovación educativa  
Instituto de Ciencias de la Educación (ICE) (2016)

ISBN: 978-84-617-5129-7

Revisión y maquetación: Neus Pellín Buades

# **Análisis de los procesos de evaluación mediante exámenes de tipo test y exámenes de desarrollo**

P. Lax Zapata<sup>1</sup>, L. Campello Blasco<sup>1</sup>, I. Ortuño Lizarán<sup>1</sup>, M.A. Noailles Gil<sup>1</sup>, G. Esquivia Sobrino<sup>1</sup>, L. Fernández Sánchez<sup>1</sup>, V. Gómez-Vicente<sup>2</sup>, N. Cuenca Navarro<sup>1</sup>, J.A. Formigós Bolea<sup>2</sup>, V. Maneu Flores<sup>2</sup>

<sup>1</sup>*Departamento de Fisiología, Genética y Microbiología*

<sup>2</sup>*Departamento de Óptica, Farmacología y Anatomía*

*Universidad de Alicante*

## **RESUMEN**

Las pruebas escritas, permiten evaluar simultáneamente múltiples objetivos de aprendizaje. Nuestra experiencia previa con cuestionarios de tipo test y de respuesta corta o de desarrollo nos hace pensar que no reflejan de la misma forma los conocimientos obtenidos por los estudiantes, es decir, que no se pueden utilizar indistintamente para evaluar las mismas competencias. Nuestro objetivo fue analizar ambos modelos de evaluación con respecto a su capacidad evaluadora en varias asignaturas de titulaciones de Ciencias de la Salud. Para ello realizamos pruebas comparadas de conocimientos adquiridos, presentando a los estudiantes distintos tipos de cuestionarios sobre una misma materia, de forma que las calificaciones de ambas partes fueran comparables. Los resultados mostraron una mayor calificación media en las pruebas de tipo test respecto a las de desarrollo (entre 0,7 y 2,6 puntos), que en todos los casos llegó a ser significativa. Entre las posibles causas de estas diferencias encontramos la experiencia previa de los alumnos en los distintos tipos de pruebas o la naturaleza de la materia evaluada, entre otras. Consideramos que sería beneficioso profundizar ampliando el estudio a diversos tipos de cuestionarios de tipo test para evaluar el potencial de los mismos.

**Palabras clave:** evaluación, test, respuesta corta, expresión escrita, cuestionarios

## 1. INTRODUCCIÓN

### 1.1 Problema y revisión de la literatura.

Uno de los asuntos fundamentales en nuestro sistema de enseñanza-aprendizaje lo constituye la evaluación. Resulta fundamental encontrar un método de evaluación idóneo que permita evaluar todos los objetivos que se plantean en cada materia, de forma fiable y válida. En este sentido, entre los distintos tipos de pruebas escritas hay grandes diferencias [Carreras i Barnés *et al.* 2009], las cuales pueden influir en gran medida en la evaluación de cada materia. Como afirman Carreras i Barnés *et al.* (2009), el sistema tradicional de evaluación mediante pruebas escritas, bien sea con preguntas de desarrollo, cortas o de tipo test, permite valorar simultáneamente multiplicidad de objetivos de aprendizaje. lo que resulta un buen procedimiento para evaluar los conocimientos y las capacidades de razonamiento y de aplicación, la capacidad de organizar y de sintetizar la información, las habilidades de escritura y la capacidad de comunicación escrita.

En estudios previos de nuestro grupo de trabajo hemos analizado distintos sistemas de evaluación aplicados a diversas asignaturas de Ciencias de la Salud [Formigós, Maneu, García-Cabanes & Palmero, 2010; García Cabanes, *et al.* 2013]. Entre los métodos de evaluación que más utilizamos para evaluar los contenidos teóricos se encuentran las pruebas de tipo test y las de respuesta corta (preguntas de respuesta abierta que solicitan del alumno la construcción de una respuesta mediante la redacción de 5 o 6 líneas). Las pruebas de tipo test, que presentan al estudiante varias respuestas preescritas con una o más alternativas válidas, son útiles para determinar si el estudiante es capaz de reconocer y discriminar la información, interpretar datos o aplicar reglas. Este tipo de cuestionarios refuerzan más el pensamiento selectivo que los procesos mentales dirigidos a la construcción del conocimiento. Las pruebas de respuesta corta en cambio, exigen una redacción y construcción de la respuesta. De esta forma permiten evaluar habilidades diferentes, como la capacidad de interpretar la evidencia o de generar una hipótesis de trabajo. Y algo muy importante, permiten valorar el uso del vocabulario y el razonamiento conceptual propio de un área de conocimiento [Carreras i Barnés *et al.*, 2009].

Por otra parte, si tenemos en cuenta la valoración, las pruebas de tipo test permiten aplicar el mismo criterio valorativo a todos los estudiantes, eliminando la posible subjetividad de las pruebas que exigen redacción de la respuesta, pero como han

indicado previamente otros autores, esto no indica que sean más objetivas respecto a la valoración del rendimiento de los estudiantes [Carreras i Barnés *et al*, 2009].

En las tablas 1 a 4 resumimos las principales fortalezas y debilidades descritos y reconocidos para ambos tipos de pruebas por otros autores y nosotros mismos [Carreras i Barnés *et al*, 2009; Lara & Cabrera, 2015].

Tabla 1. Principales fortalezas de las pruebas de tipo test.

- Permiten aplicar el mismo criterio valorativo a todos los estudiantes, eliminando la posible subjetividad de las pruebas que exigen redacción de la respuesta.
- Permiten valorar en una única prueba la adquisición de conocimientos referentes a múltiples temas diferentes o un mismo tema en profundidad.
- Con el tiempo se puede disponer de una batería de preguntas que se pueden poner a disposición de los estudiantes a modo de pruebas de autoevaluación y/o de *feedback* o retroalimentación para el profesor. Si esta batería es suficientemente grande podría formar parte de las pruebas de evaluación, utilizando un sistema de selección aleatorio.

Tabla 2. Principales debilidades de las pruebas de tipo test.

- El hecho de que permitan aplicar el mismo criterio valorativo a los distintos estudiantes no indica que sean más objetivas respecto a la valoración del rendimiento de los mismos.
- Consideramos que la elaboración de un buen cuestionario con preguntas de tipo test es muy complicada.
- No permite la valoración de competencias que impliquen construcción, estructuración y exposición de conocimiento, no estimula el uso de vocabulario propio del área de conocimiento.

Tabla 3. Principales fortalezas de las pruebas de respuesta corta.

- Dan la sensación (al profesorado) de ser más objetivas respecto a la valoración del rendimiento de los estudiantes.
- Permiten mayor libertad al estudiante para elaborar la respuesta.
- Permiten valorar y fomentan el uso del vocabulario y el razonamiento

conceptual propio de un área de conocimiento.

- Pueden utilizarse en diversas actividades, como en "preguntas del minuto" al final de una clase presencial.

Tabla 4. Principales debilidades de las pruebas de respuesta corta.

- Hay que asegurarse de aplicar el mismo criterio valorativo a todos los estudiantes.
- No se prestan a la obtención de una batería de preguntas de autoevaluación o retroalimentación rápida puesto que requieren la corrección por parte del profesor.

Preocupados desde hace tiempo por esta cuestión, hemos realizados varios análisis previos y nuestra experiencia nos muestra que en nuestras asignaturas los cuestionarios de tipo test y los de respuesta corta no reflejan de la misma forma las competencias adquiridas por los estudiantes. Estos muestran una mayor dificultad para construir y redactar una respuesta que para reconocer una respuesta correcta entre varias preescritas. En los últimos años venimos observando un deterioro en la expresión escrita de los estudiantes, que se manifiesta en las pruebas que exigen redacción y que afecta a la calificación final de las mismas. Así, hemos comprobado que evaluando mediante pruebas de tipo test y mediante pruebas de respuesta corta, los estudiantes obtienen calificaciones distintas [Campello Blasco *et al.*, 2016]. En nuestros estudios previos hemos comparado calificaciones obtenidas con los dos tipos de cuestionarios aplicados a las mismas asignaturas pero, al no estar diseñados al efecto, las cuestiones no incidían exactamente sobre los mismos problemas, sino sobre puntos distintos de la misma asignatura.

### 1.3 Propósito.

Nuestro objetivo principal fue analizar ambos modelos de evaluación, cuestionarios de tipo test y cuestionarios de respuestas breves, con respecto a su capacidad de evaluar las competencias alcanzadas por los estudiantes en dos asignaturas de Ciencias de la Salud impartidas por profesores de nuestro grupo de trabajo, mediante el diseño de cuestionarios específicos creados al efecto.

## 2. METODOLOGÍA

### 2.1. Descripción del contexto y de los participantes

La red está formada por un equipo multidisciplinar (Tabla 5), con docencia en distintas asignaturas pertenecientes a diversas áreas de conocimiento.

En este trabajo se centró en las asignaturas: "Fisiología" (asignatura básica, 1er curso, Grado en Nutrición Humana y Dietética) y "Farmacología" (asignatura obligatoria, 2º curso, Grado en Óptica y Optometría). Los cuestionarios se realizaron durante el curso académico 2015-2016.

Tabla 5. Miembros de la red y filiación.

Miembro de la red	Departamento (Universidad)
Victoria Maneu Juan Formigós M. Violeta Gómez Vicente	Óptica, Farmacología y Anatomía (Universidad de Alicante)
Nicolás Cuenca Pedro Lax Laura Campello Laura Fernández Gema Esquiva M. Agustina Noailles Isabel Ortuño Lizarán	Fisiología, Genética y Microbiología (Universidad de Alicante)

### 2.2. Materiales e instrumentos

Se diseñaron pruebas formadas por dos cuestionarios cada una, uno de tipo test y otro de preguntas de respuesta abierta corta, de forma que se preguntara sobre el mismo contenido y específicamente sobre el mismo problema o cuestión. De esta forma pretendíamos que las calificaciones en ambos cuestionarios fueran comparables entre sí.

En la asignatura "Farmacología" los cuestionarios estaban compuestos por un total de entre 5 y 8 preguntas de respuesta corta y entre 5 y 9 preguntas de tipo test.

En la asignatura "Fisiología", el cuestionario se compuso de 6 preguntas de respuesta corta y 9 preguntas de tipo test.

### 2.3. Procedimientos

Los dos modelos de cuestionarios se plantearon en la misma sesión. En primer lugar se presentó la prueba de desarrollo y a continuación, una vez entregada esta primera parte, se presentó la prueba de tipo test, para evitar que la lectura de las respuestas del test influyera en la redacción de las respuestas de los estudiantes.

En la asignatura "Fisiología" se realizó una prueba única, mientras que en la asignatura "Farmacología" se realizaron tres pruebas similares, para evaluar el posible efecto del "entrenamiento" en los estudiantes.

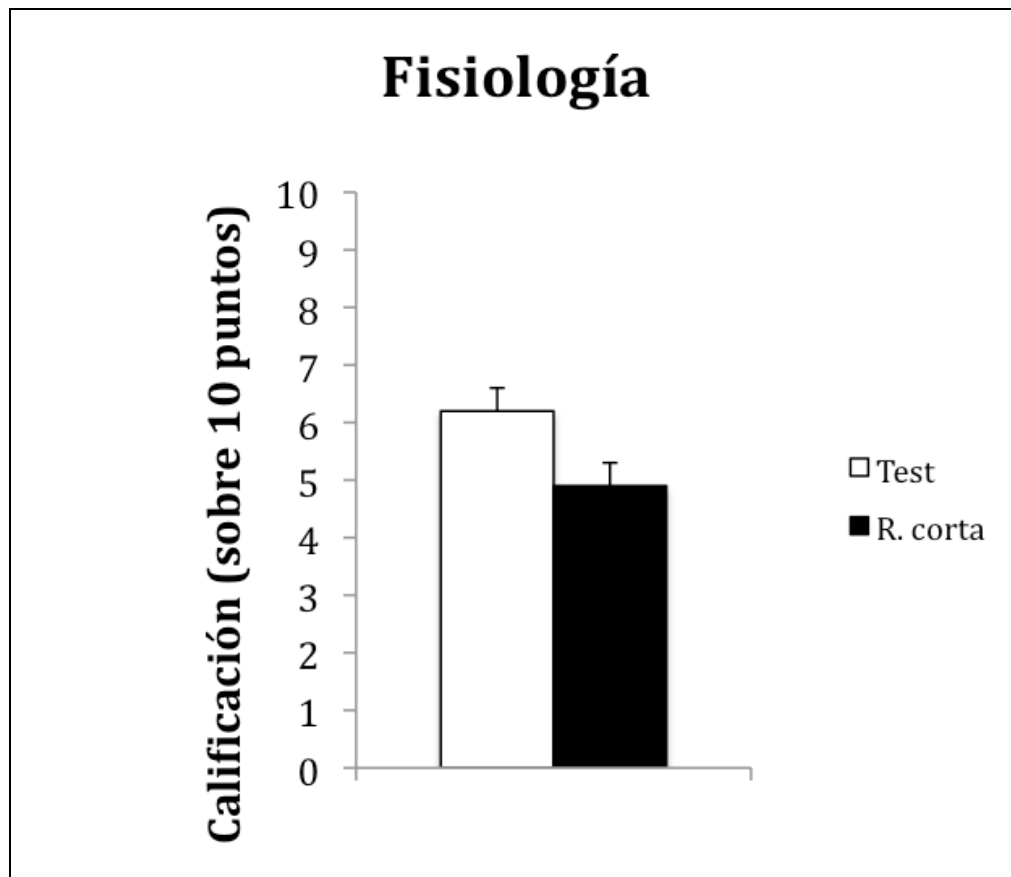
Los resultados se analizaron mediante comparación de la media de las calificaciones de cada una de las pruebas. Se analizaron entre 17 y 92 resultados en cada una de las pruebas. Los datos se presentan como la media  $\pm$  la desviación estándar, siempre respecto a 10 puntos de calificación máxima. Los resultados de cada una de las modalidades estudiadas se analizaron mediante la prueba t de Student. Se consideraron significativos valores de  $p < 0,05$ .

## 3. RESULTADOS

En la prueba realizada en la asignatura "Fisiología", los resultados fueron de  $6,2 \pm 0,4$  puntos en el cuestionario de tipo test y de  $4,9 \pm 0,4$  puntos en parte de la prueba correspondiente a las preguntas de desarrollo. La diferencia entre las dos modalidades resultó estadísticamente significativa (Figura 1).

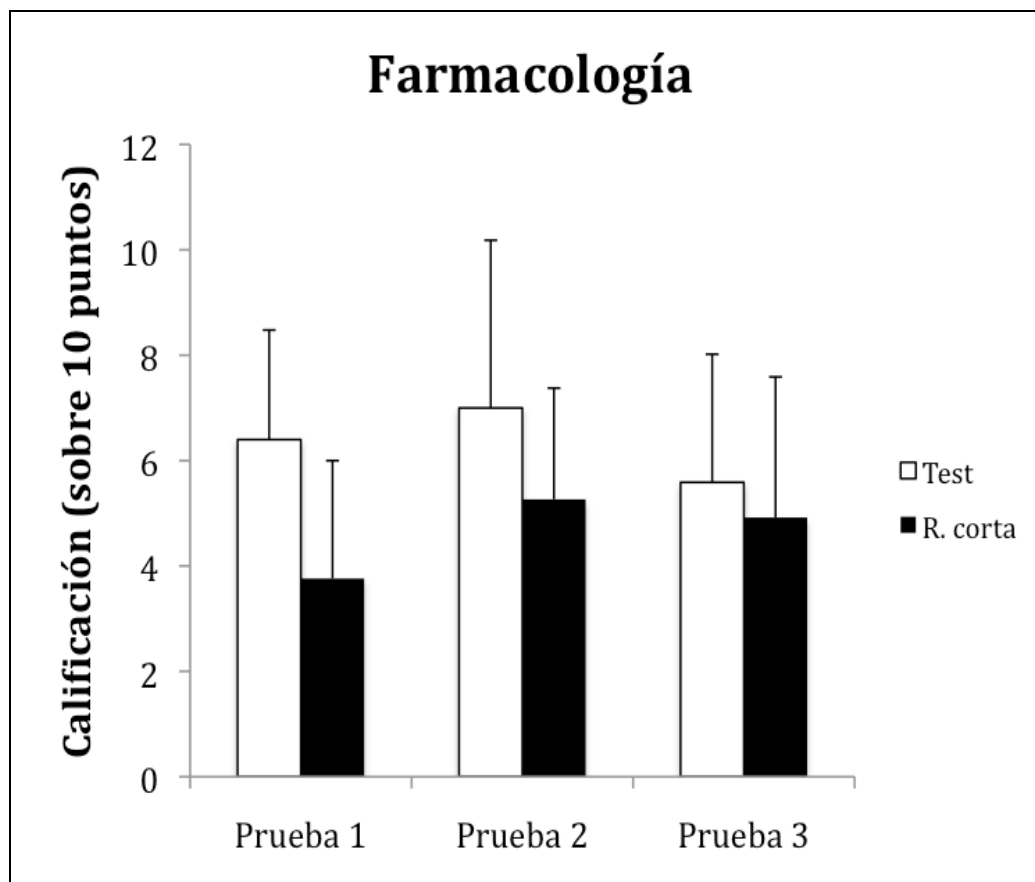


Figura 1. Calificaciones (sobre 10 puntos) obtenidas en la prueba de tipo test y de respuesta corta en la asignatura Fisiología, del Grado en Nutrición Humana y Dietética. La prueba se realizó durante el curso 2015-16.



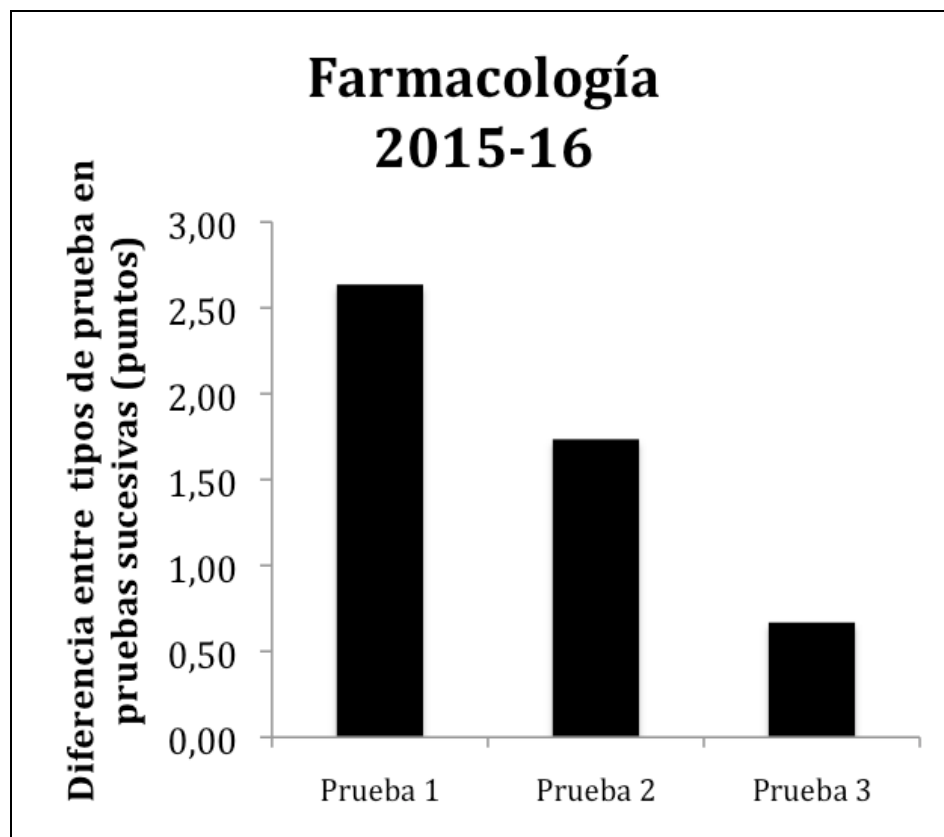
En la asignatura "Farmacología" se realizaron un total de tres pruebas. Las calificaciones medias de los cuestionarios de tipo test fueron de  $6,4 \pm 2,1$ ;  $7,0 \pm 3,2$  y  $5,6 \pm 2,4$  puntos, mientras que en las preguntas de desarrollo las calificaciones fueron de  $3,8 \pm 2,2$ ;  $5,3 \pm 2,1$  y  $4,9 \pm 2,7$  puntos respectivamente (Figura 2).

Figura 2. Calificaciones (sobre 10 puntos) obtenidas en pruebas de tipo test y de respuesta corta en la asignatura "Farmacología", del Grado en Óptica y Optometría. Las pruebas se realizaron en el curso 2015-16.



La realización de varias pruebas en el mismo curso nos permitió también evaluar el posible efecto del entrenamiento en las pruebas. Si bien en todos los casos las notas fueron superiores en las preguntas de tipo test, la diferencia se fue acortando, pasando de 2,6 puntos en la primera prueba a 1,7 en la segunda y a 0,7 en la tercera (Figura 3). En todos los casos las diferencias entre las dos modalidades fueron estadísticamente significativas.

Figura 3. Diferencia entre las calificaciones (sobre 10 puntos) obtenidas en las pruebas de tipo test y pruebas de respuesta corta en la asignatura "Farmacología", del Grado en Óptica y Optometría. Las pruebas se realizaron en el curso 2015-16.



Sobre la opinión de los profesores implicados en este estudio, nuestra percepción es que mediante los cuestionarios de respuesta corta evaluamos mejor las competencias de los estudiantes, puesto que valoramos también la expresión, la coherencia y el hilo de los razonamientos. Esta idea se ha visto reforzada en los últimos años, al constatar la diferencia en las calificaciones de ambos cuestionarios y comparar las respuestas de los mismos alumnos en las pruebas de uno y otro formato.

#### 4. CONCLUSIONES

Las pruebas realizadas en el contexto de esta red, mostraron que en todos los casos las calificaciones fueron superiores (en mayor o menor grado) en las pruebas de tipo test respecto a las preguntas que exigían al estudiante una construcción de la respuesta.

La realización de varias pruebas durante el mismo curso en la asignatura "Farmacología", sí se redujo la diferencia entre los dos tipos de prueba. Este aparente efecto del entrenamiento se trata de un resultado preliminar que debería confirmarse realizando más pruebas en diversas asignaturas y cursos. Aún siendo preliminar, consideramos que esta prueba es muy interesante, pues demostraría una cuestión que observamos los profesores implicados en este trabajo, y es que nosotros sí apreciamos una mejora en la expresión cuando los mismos estudiantes se someten repetidamente a pruebas o actividades que requieren el ejercicio de la expresión escrita. La mejora en la expresión va también acompañada de una mejora en la utilización de lenguaje específico del área de conocimiento. Por esto recomendamos actividades que la fomenten para mejorar las habilidades de los estudiantes. En los últimos años hemos observado un deterioro progresivo de la capacidad de redacción y construcción de respuestas por parte de los estudiantes, por lo que nos parece no sólo conveniente, sino necesario mejorar la expresión escrita. Los mismos estudiantes nos han manifestado abiertamente su dificultad en la redacción, hecho que achacamos fundamentalmente a la falta de práctica debido al uso cada vez más extendido y generalizado (así manifestado por los alumnos) de la realización de pruebas de tipo test.

El cambio en el proceso evaluador del proceso de convergencia en el Espacio Europeo de Educación Superior (EEES) se ha sumado a los cambios evidentes en el modelo enseñanza-aprendizaje. Las competencias adquiridas por los estudiantes al final de su formación implican el uso integrado de conocimientos, habilidades y actitudes. La expresión y la redacción son necesarias para el desarrollo profesional completo, para expresarse y comunicarse con los colegas de profesión y, en caso de titulaciones de Ciencias de la Salud, todavía más, puesto que se exige un trato de comunicación clara, directa y sencilla con los pacientes, que resulta fundamental en muchos aspectos del ejercicio profesional. La evaluación debe pues valorar estas habilidades.

La American Association for Higher Education, la evaluación es más eficaz cuando el aprendizaje se entiende como multidimensional. El aprendizaje implica no sólo lo que saben los estudiantes, sino lo que pueden hacer con lo que saben [AAHE, 1992]. Además, la evaluación en sí misma puede ser un acto formativo y constituir una oportunidad de aprendizaje, que promueva la consecución de las competencias por todos los estudiantes [Cano García, 2008; Dunn & Mulvenon, 2009; Scriven, 1967].

Según nuestro criterio, las habilidades de redacción deben trabajarse más en nuestro contexto y su inclusión en las pruebas de evaluación es una de las vías para fomentarlo.

Además de las ventajas e inconvenientes que presentan los dos modelos de evaluación discutidos en este trabajo, queremos resaltar que ambos presentan una serie de debilidades. Tienen una capacidad limitada de demostrar la asimilación de conceptos entre distintas unidades de la materia, lo cual sí que se conseguiría con preguntas que permitieran un desarrollo mayor.

Destacamos pues de nuestro trabajo, por una parte, que las pruebas de tipo test y las de respuesta corta no se muestran igualmente válidas para evaluar las distintas capacidades. Consideramos que es necesario hacer un análisis concienzudo de cuáles son las pruebas de evaluación óptimas en cada asignatura según los objetivos que se deseen alcanzar y elegir las idóneas en cada caso. Según nuestra opinión, se deberían incluir más actividades que potenciaran la habilidad de la expresión escrita de los estudiantes e incluir preguntas de respuesta corta o abierta en las pruebas escritas para potenciar su desarrollo.

## **5. DIFICULTADES ENCONTRADAS**

Para obtener resultados extrapolables decidimos centrarnos en asignaturas con cierta similitud en cuanto a los requerimientos teórico-prácticos, incluyendo las asignaturas de Fisiología y Farmacología. Pero hay que tener en cuenta que para poder extrapolar realmente las conclusiones, se debería realizar un estudio más amplio, incluyendo diversas asignaturas y también con distinta carga teórica y práctica.

Una de las cuestiones que tuvimos que decidir fue el tipo de cuestionarios de tipo test que incluimos en nuestro estudio. Decidimos utilizar un modelo similar al que utilizamos en cada una de nuestras asignaturas: en la asignatura "Farmacología" se presentaron cuestiones con tres alternativas de respuesta, que incluían una única válida. En la asignatura "Fisiología" se presentaron cuestiones con cinco alternativas de respuesta, que incluían una única válida. Nos planteamos la conveniencia de utilizar otros modelos, si bien finalmente nos decidimos por estos por ser los que habitualmente utilizamos en estas dos asignaturas. No obstante, como expresamos en el punto siguiente, consideramos que sería útil realizar el estudio con distintas opciones de respuestas válidas.

## **6. PROPUESTAS DE MEJORA**

Como hemos desarrollado en la discusión, en este trabajo presentamos datos que confirman nuestra idea de que las pruebas de tipo test fallan a la hora de evaluar determinadas competencias que se deben exigir a los estudiantes de las titulaciones implicadas en el estudio. Pero también parece evidente que las pruebas de tipo test merecen un estudio más detallado. Creemos que se deberían probar y comparar diversas modalidades (por ejemplo cambiando/ampliando el número de respuestas posibles o incluyendo la posibilidad de más de una respuesta válida, entre otras), para encontrar las pruebas que reflejen de forma más real los conocimientos de los estudiantes en las asignaturas con las que trabajamos. Confiamos en que una optimización de estas pruebas debería equipararla con las pruebas de respuesta redactada al menos para evaluar algunas de las competencias que se persiguen.

Como comentamos en un trabajo previo [Campello Blasco *et al.*, 2016], no debemos descartar cualquier tipo de responsabilidad en la formulación de las pruebas, y creemos que sería útil proponer un curso de formación para todo el profesorado sobre el diseño de cuestionarios de tipo test, lo cual podría contribuir a mejorar los cuestionarios. A pesar de estas consideraciones, los autores de este trabajo mantenemos la idea de que existe una necesidad de que los titulados adquieran las destrezas necesarias de comunicación y expresión que les permitan expresarse correctamente, utilizando el lenguaje profesional adecuado. Según nuestra opinión, esto pasa por dar un peso mayor a los exámenes que exigen un esfuerzo de redacción.

## **7. PREVISIÓN DE CONTINUIDAD**

Los autores de este trabajo consideramos que es esencial encontrar sistemas de calificación que reflejen de forma objetiva y precisa los conocimientos y competencias adquiridas por los estudiantes. Creemos que es necesario seguir trabajando en este tema para conseguir diseñar las pruebas idóneas. La exploración de las diversas tipologías de pruebas de tipo test puede ser una buena vía de continuación de este trabajo en futuras ediciones.

También consideramos interesante profundizar en la metodología de elaboración de cuestionarios de desarrollo. En este sentido hemos encontrado alguna bibliografía al respecto. Una búsqueda intensiva y un grupo de estudio y trabajo, quizá con la

realización de algún taller con la ayuda del ICE de la Universidad de Alicante, podría ser beneficioso para todos.

## 8. REFERENCIAS BIBLIOGRÁFICAS

- American Association for Higher Education (AAHE). *Nine Principles of Good Practice for Assessing Student Learning*, 1992. Washington, DC: U.S. Recuperado de <http://condor.depaul.edu/tla/Learning/Student%20Learning%20AAHE.pdf>
- Campello Blasco, L., Esquiva Sobrino G, Noailles Gil, M.A., Fernández Sánchez, L., Gómez-Vicente, M.V., Cuenca Navarro, N., Formigós Bolea, J., Maneu Flores, V., Lax Zapata, P. (2016). Análisis de los procesos de la evaluación en las nuevas titulaciones de grado. En *Innovaciones Metodológicas en Docencia Universitaria: Resultados de Investigación*. J.D. Álvarez Teruel, S. Grau Company y M.T. Tortosa Ybáñez (Coords.) Universidad de Alicante.
- Cano García, M.E. (2008) La evaluación por competencias en la educación superior. *Revista de curriculum y formación del profesorado*. Recuperado de <https://www.ugr.es/~recfpro/rev123COL1.pdf>
- Carreras i Barnés, J., Branda, L.A., Castro, A., Fenoll, M.L., Gual, A., Mahy J.N., Martínez J.M., Nolla, M., Palés, J., Pérez, J., Portero, M. y Romá, J. (2009) *Guía para la evaluación de competencias en Medicina*. Agència per a la Qualitat del Sistema Universitari de Catalunya.
- Dunn, K.E and Mulvenon, S.W. (2009). A Critical Review of Research on Formative Assessments: The Limited Scientific Evidence of the Impact of Formative Assessments in Education. En *Practical Assessment Research & Evaluation*, 14 (7). Recuperado de <http://pareonline.net/getvn.asp?v=14&n=7>
- Formigós Bolea, J.A., Maneu Flores, V., García Cabanes, MC, Palmero, M.M. (2010) Experiencia de evaluación del aprendizaje de Farmacología en la Universidad de Alicante. En *Evaluación de los aprendizajes en el Espacio Europeo de Educación Superior*. Salvador Grau Company, Cecilia Gómez Lucas (Coord.). Alcoy : Marfil, Alicante, Universidad de Alicante.
- García Cabanes, MC., Formigós Bolea, J.A., Lax Zapata, P., Hurtado Sánchez, J.A., Campello Blasco, L., Gómez Vicente, M.V., López Rodríguez, D., Esquiva Sobrino, G., Cuenca Navarro, N., Maneu Flores, V. (2013). Análisis de diversos métodos de evaluación implantados en distintas asignaturas de los títulos de

grado. En M.T. Tortosa Ybáñez, J.D. Álvarez Teruel, N. Pellín Buades (coords) [Recurso electrónico]: *XI Jornadas de Redes de Investigación en Docencia Universitaria: Retos de futuro en la enseñanza superior: docencia e investigación para alcanzar la excelencia académica* (pp. 1886-1896) Alicante: Universidad de Alicante.

Lara, F. y Cabrera, M. (2015) *Fichas de procedimientos de evaluación educativa*. Unidad de Gestión Curricular. Universidad de las Américas (UDLA).

Scriven, M. (1967). The methodology of evaluation. En R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39-83). Chicago: Rand-McNally.